

## SEQ SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      October 20, 2003, 19:02:39 ; Search time 4689 Seconds
              (without alignments)
              16594.180 Million cell updates/sec
```

```
Title:          US-09-824-551-1
Perfect score:  1902
Sequence:       1 ggtaggagtaaaaaacgcag.....ggccccggcgtccaaggaac 1902
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :                               GenEmbl:*
1:      gb_ba:*
2:      gb_htg:*
3:      gb_in:*
4:      gb_om:*
5:      gb_ov:*
6:      gb_pat:*
7:      gb_ph:*
8:      gb_pl:*
9:      gb_pr:*
10:     gb_ro:*
11:     gb_sts:*
12:     gb_sy:*
13:     gb_un:*
14:     gb_vi:*
15:     em_ba:*
16:     em_fun:*
17:     em_hum:*
18:     em_in:*
19:     em_mu:*
20:     em_om:*
21:     em_or:*
22:     em_ov:*
23:     em_pat:*
24:     em_ph:*
25:     em_pl:*
26:     em_ro:*
27:     em_sts:*
```

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Match	Query Length	DB	ID	Description
	1	1902	100.0	1902	6	AX397915 <i>wo 02/20799 Rathe</i>	AX397915 Sequence
	2	1902	100.0	1902	6	AX398211 <i>wo 02/20793 "</i>	AX398211 Sequence
c	3	1902	100.0	309400	6	AX127153 <i>EP 1108790 Nakagawa</i>	AX127153 Sequence
c	4	1902	100.0	325651	1	AP005283 <i>GenBank "</i>	AP005283 Corynebac
	5	1311	68.9	1311	6	AX123323 <i>EP</i>	AX123323 Sequence
	6	1311	68.9	1311	6	BD165440 <i>JP - Nakagawa</i>	BD165440 Novel pol
c	7	773.2	40.7	302070	1	AP005223	AP005223 Corynebac
c	8	639	33.6	639	6	AX123321	AX123321 Sequence
c	9	639	33.6	639	6	BD165438	BD165438 Novel pol
	10	404	21.2	1052	6	AX377471	AX377471 Sequence
c	11	319	16.8	588	6	AX123322	AX123322 Sequence
c	12	319	16.8	588	6	BD165439	BD165439 Novel pol
	13	291	15.3	759	6	AX065953 <i>Pompejus wo 01/00842</i>	AX065953 Sequence
	14	191	10.0	636	6	AX123320	AX123320 Sequence
	15	191	10.0	636	6	BD165437	BD165437 Novel pol
	16	108	5.7	339650	1	SCO939108	AL939108 Streptomy
	17	96.4	5.1	1149	6	AX122285	AX122285 Sequence
	18	96.4	5.1	1149	6	BD164402	BD164402 Novel pol
c	19	96.4	5.1	337200	1	AP005280	AP005280 Corynebac
c	20	96.4	5.1	349980	6	AX127149	AX127149 Sequence
c	21	96.4	5.1	349980	6	AX127150	AX127150 Sequence
	22	76.6	4.0	299050	1	SCO939119	AL939119 Streptomy
	23	73.2	3.8	2760	1	SRSENRS	Y14336 Streptomyce
c	24	70.8	3.7	298450	1	SCO939107	AL939107 Streptomy
	25	70.4	3.7	10035	1	AE001999	AE001999 Deinococc
c	26	63.8	3.4	291000	1	SCO939105	AL939105 Streptomy
c	27	61.6	3.2	255447	2	AC096032	AC096032 Rattus no
	28	61	3.2	1965	1	BSU18135	U18135 Pseudomonas
c	29	61	3.2	196050	1	AL646058	AL646058 Ralstonia
	30	60.6	3.2	632	6	AX435920	AX435920 Sequence
	31	60.6	3.2	299925	1	AP005043	AP005043 Streptomy
	32	60.4	3.2	753	6	AX065957	AX065957 Sequence
	33	60.4	3.2	1032	6	AX376724	AX376724 Sequence

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 18:19:14 ; Search time 385 Seconds  
(without alignments)  
13335.925 Million cell updates/sec

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcgtccaaggaac 1902

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	1902	100.0	1902	24	ABL40476 <i>wo 02/20799</i>	C. glutamicum luxS
c	2	1902	100.0	309400	22	AAH68534 <i>ET 1108710</i>	C glutamicum codin
	3	1311	68.9	1311	22	AAH68204	C glutamicum codin
	4	858	45.1	858	25	ACA01369	C. glutamicum deri
c	5	639	33.6	639	22	AAH68202	C glutamicum codin
	6	415	21.8	426	25	ACA01379	C. glutamicum deri
	7	404	21.2	1052	24	ABA91930	Corynebacterium gl
c	8	319	16.8	588	22	AAH68203	C glutamicum codin
	9	291	15.3	759	22	AAF71286	Corynebacterium gl
c	10	280	14.7	549	25	ACA01390	C. glutamicum deri
	11	191	10.0	636	22	AAH68201	C glutamicum codin
	12	96.4	5.1	1149	22	AAH67166	C glutamicum codin
c	13	96.4	5.1	349980	22	AAH68530	C glutamicum codin
c	14	96.4	5.1	349980	22	AAH68531	C glutamicum codin
	15	60.6	3.2	632	24	ABK77044	Bacillus lichenifo
	16	60.4	3.2	753	22	AAF71288	Corynebacterium gl
	17	60.4	3.2	1032	24	ABA95154	C. glutamicum lysR
	18	60	3.2	630	22	AAH67165	C glutamicum codin
	19	58.2	3.1	2538	20	AAX21398	Central fragment o
	20	58.2	3.1	155074	24	ABN85735	Human genomic regi
	21	56.2	3.0	12531	23	AAS59530	Propionibacterium
	22	53.4	2.8	233	24	ABN23371	Human ORFX polynuc
	23	53	2.8	353	24	ABA91931	Corynebacterium gl
c	24	52.6	2.8	5944	23	AAS59604	Propionibacterium
c	25	52	2.7	25360	22	AAF88314	S. spinosa DNA fra
c	26	52	2.7	29736	22	AAF88317	S. spinosa DNA fra
	27	51.6	2.7	601	24	ABK77051	Bacillus lichenifo
	28	51	2.7	840	22	AAS63305	DNA encoding Esche
c	29	50	2.6	1236	24	AAD44410	Human huntington (
c	30	50	2.6	1662	23	AAS86725	DNA encoding novel
c	31	50	2.6	3177	23	AAS93207	DNA encoding novel
c	32	50	2.6	3177	23	AAS94065	DNA encoding novel
	33	48	2.5	1058	24	ABK76966	Bacillus lichenifo
c	34	48	2.5	2422	23	AAS85331	DNA encoding novel
c	35	48	2.5	4629	23	ABL29279	Drosophila melanog
c	36	48	2.5	6693	23	ABL29278	Drosophila melanog
c	37	46.8	2.5	486	22	AAF75507	Polyglutamine trac
	38	46.8	2.5	1122	23	AAS93196	DNA encoding novel
c	39	46.8	2.5	3297	23	AAS91437	DNA encoding novel
	40	46.6	2.5	642	19	AAV39839	Pseudomonas fluore
	41	46.6	2.5	642	19	AAV31652	Nucleotide sequenc
	42	46.6	2.5	642	20	AAX99370	P. fluorescens mod
c	43	46.6	2.5	2109	23	AAS81477	DNA encoding novel
c	44	46.4	2.4	1635	22	ABA49946	Human breast cell
c	45	46.4	2.4	1635	22	ABA67865	Human foetal liver

# ALIGNMENTS

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 20:40:15 ; Search time 376 Seconds  
(without alignments)  
13284.296 Million cell updates/sec

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcggtccaaggaac 1902

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result			%		DB	ID	Description
	No.	Score	Match	Length			
	1	1902	100.0	1902	10	US-09-824-551-1 <i>HESS</i>	Sequence 1, Appli
c	2	1902	100.0	3309400	10	US-09-738-626-1 <i>Not a sequence</i>	Sequence 1, Appli
	3	1311	68.9	1311	10	US-09-738-626-3239	Sequence 3239, Ap
c	4	639	33.6	639	10	US-09-738-626-3237	Sequence 3237, Ap
c	5	319	16.8	588	10	US-09-738-626-3238	Sequence 3238, Ap
	6	191	10.0	636	10	US-09-738-626-3236	Sequence 3236, Ap
	7	96.4	5.1	1149	10	US-09-738-626-2201	Sequence 2201, Ap
	8	60.6	3.2	632	10	US-09-974-300-4335	Sequence 4335, Ap
	9	60.6	3.2	717	14	US-10-156-761-5606	Sequence 5606, Ap
	10	60.6	3.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
	11	60.4	3.2	1032	11	US-09-867-537A-1	Sequence 1, Appli
	12	60	3.2	630	10	US-09-738-626-2200	Sequence 2200, Ap
	13	58.2	3.1	155074	13	US-10-026-188-6	Sequence 6, Appli
	14	57.2	3.0	639	14	US-10-156-761-2864	Sequence 2864, Ap
c	15	57.2	3.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
	16	55.6	2.9	663	14	US-10-156-761-1453	Sequence 1453, Ap
	17	55.2	2.9	639	14	US-10-156-761-6454	Sequence 6454, Ap
	18	53.6	2.8	681	14	US-10-156-761-6016	Sequence 6016, Ap
	19	53.2	2.8	669	14	US-10-156-761-1976	Sequence 1976, Ap
	20	51.6	2.7	601	10	US-09-974-300-4342	Sequence 4342, Ap
c	21	50	2.6	1236	14	US-10-077-584-3	Sequence 3, Appli
	22	48.2	2.5	669	14	US-10-156-761-4665	Sequence 4665, Ap
	23	48	2.5	1058	10	US-09-974-300-4257	Sequence 4257, Ap
	24	47.2	2.5	744	14	US-10-156-761-5052	Sequence 5052, Ap
	25	47	2.5	648	14	US-10-156-761-4438	Sequence 4438, Ap
c	26	47	2.5	1972	12	US-10-017-161-1981	Sequence 1981, Ap
	27	46.6	2.5	672	14	US-10-156-761-4521	Sequence 4521, Ap
c	28	46.4	2.4	1635	9	US-09-864-761-20241	Sequence 20241, A
c	29	46.4	2.4	1973	9	US-09-864-761-3471	Sequence 3471, Ap
	30	46.2	2.4	669	14	US-10-156-761-7370	Sequence 7370, Ap
	31	45.8	2.4	38918	12	US-10-017-161-2049	Sequence 2049, Ap
	32	45.6	2.4	672	14	US-10-156-761-4519	Sequence 4519, Ap
	33	45.6	2.4	1075	9	US-09-864-761-19241	Sequence 19241, A
	34	45.6	2.4	1403	9	US-09-864-761-2513	Sequence 2513, Ap
c	35	45.4	2.4	88421	10	US-09-976-059-1	Sequence 1, Appli
	36	45	2.4	612	14	US-10-156-761-3119	Sequence 3119, Ap
c	37	45	2.4	3367	12	US-09-890-688-129	Sequence 129, App
c	38	45	2.4	5085	14	US-10-198-846-9854	Sequence 9854, Ap
	39	45	2.4	5499	14	US-10-198-846-12674	Sequence 12674, A
	40	44.4	2.3	446	9	US-09-864-761-20699	Sequence 20699, A
	41	43.8	2.3	1104	9	US-09-815-242-6579	Sequence 6579, Ap
	42	43.8	2.3	10391	10	US-09-070-927A-133	Sequence 133, App
	43	43	2.3	729	14	US-10-156-761-3236	Sequence 3236, Ap
	44	42.8	2.3	720	9	US-09-815-242-9828	Sequence 9828, Ap
	45	42.8	2.3	1234	12	US-10-182-504-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-824-551-1

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      October 20, 2003, 19:21:29 ; Search time 2764 Seconds
              (without alignments)
              16724.715 Million cell updates/sec
```

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcgtccaaggaac 1902

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*
```

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
c	1	57	3.0	644	12	BM605179	BM605179 170006870
	2	56.6	3.0	712	13	BX416727	BX416727 BX416727
	3	55	2.9	885	13	BX425603	BX425603 BX425603
	4	54.8	2.9	895	29	CNS0071A	AL066286 Drosophil
	5	50.8	2.7	884	29	CNS006U0	AL065923 Drosophil
c	6	50.6	2.7	818	29	BZ229437	BZ229437 CH230-401
c	7	50.2	2.6	570	14	CD344209	CD344209 EtESTef78
c	8	49.4	2.6	638	13	BW260725	BW260725 BW260725
c	9	49.4	2.6	712	28	BH898573	BH898573 MB61p8H7
	10	49.2	2.6	711	28	BZ098560	BZ098560 CH230-235
	11	49.2	2.6	922	29	CNS0073W	AL066784 Drosophil
c	12	48.8	2.6	1005	29	CNS02LA1	AL202546 Tetraodon
	13	48.8	2.6	1132	29	BZ559201	BZ559201 pacs2-164
c	14	48.4	2.5	711	12	BM588321	BM588321 170006873
	15	48.2	2.5	439	12	BJ348296	BJ348296 BJ348296
	16	48.2	2.5	468	12	BJ348080	BJ348080 BJ348080
	17	48.2	2.5	505	12	BJ406583	BJ406583 BJ406583
	18	48.2	2.5	522	12	BJ353136	BJ353136 BJ353136
	19	48.2	2.5	555	12	BJ403974	BJ403974 BJ403974
	20	48.2	2.5	695	12	BJ439038	BJ439038 BJ439038
c	21	48	2.5	664	12	BM595824	BM595824 170006874
c	22	47.8	2.5	540	9	AW464269	AW464269 BP230015A
	23	47.6	2.5	689	12	BJ485119	BJ485119 BJ485119
	24	47.2	2.5	902	29	CNS006QP	AL065804 Drosophil
	25	47.2	2.5	1201	13	BX356664	BX356664 BX356664
c	26	47	2.5	789	28	BZ115869	BZ115869 CH230-255
c	27	46.6	2.5	494	12	BM641696	BM641696 170006873
c	28	46.4	2.4	450	14	CA915472	CA915472 PCS01853X
c	29	46.4	2.4	481	14	CA915480	CA915480 PCS04158
c	30	46.4	2.4	506	14	CA899712	CA899712 PCEP01447
c	31	46.4	2.4	512	14	CA915525	CA915525 PCSC21474
c	32	46.4	2.4	598	14	CA915510	CA915510 PCSC17379
c	33	46.4	2.4	617	13	BW314712	BW314712 BW314712
c	34	46.4	2.4	638	12	BM618697	BM618697 170006874
c	35	46.4	2.4	699	12	BM630395	BM630395 170006875
c	36	46.2	2.4	796	29	BZ556455	BZ556455 pacs1-60
c	37	46	2.4	572	29	CNS043SN	AL273200 Tetraodon
c	38	46	2.4	602	13	BW192111	BW192111 BW192111
c	39	46	2.4	603	14	CB555832	CB555832 MMSP0006
c	40	46	2.4	621	13	BW265801	BW265801 BW265801
c	41	46	2.4	658	13	BW265514	BW265514 BW265514
c	42	46	2.4	663	13	BW259047	BW259047 BW259047
	43	45.8	2.4	477	14	CA743784	CA743784 wrils.pk0
	44	45.8	2.4	585	13	BQ190440	BQ190440 UI-R-DN1-
c	45	45.6	2.4	560	14	CD569824	CD569824 EtESTeg05



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 19:23:55 ; Search time 100 Seconds  
(without alignments)  
8395.107 Million cell updates/sec

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcggtccaaggaac 1902

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	60.4	3.2	1032	4	US-09-897-537A-1	Sequence 1, Appli
	2	58.2	3.1	2588	2	US-08-796-414B-6	Sequence 6, Appli
	3	51	2.7	840	4	US-09-651-656-106	Sequence 106, App
	4	51	2.7	840	4	US-09-650-855-106	Sequence 106, App
	5	48.8	2.6	678	4	US-09-252-991A-794	Sequence 794, App
c	6	48.8	2.6	867	4	US-09-252-991A-848	Sequence 848, App
c	7	48.8	2.6	2577	4	US-09-252-991A-828	Sequence 828, App
	8	48.4	2.5	810	4	US-09-252-991A-3647	Sequence 3647, Ap
	9	48.4	2.5	879	4	US-09-252-991A-3602	Sequence 3602, Ap
c	10	48.4	2.5	1632	4	US-09-252-991A-3859	Sequence 3859, Ap
	11	46.6	2.5	642	1	US-08-761-258-4	Sequence 4, Appli



29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%					Description
	No.	Score	Match	Length	ID	
	1	1902	100.0	1902	6	AX397915 Sequence
	2	1902	100.0	1902	6	AX398211 Sequence
c	3	1902	100.0	309400	6	AX127153 Sequence
c	4	1902	100.0	325651	1	AP005283 Corynebac
	5	1311	68.9	1311	6	AX123323 Sequence
	6	1311	68.9	1311	6	BD165440 Novel pol
c	7	639	33.6	639	6	AX123321 Sequence
c	8	639	33.6	639	6	BD165438 Novel pol
	9	404	21.2	1052	6	AX377471 Sequence
c	10	319	16.8	588	6	AX123322 Sequence
c	11	319	16.8	588	6	BD165439 Novel pol
	12	291	15.3	759	6	AX065953 Sequence
	13	191	10.0	636	6	AX123320 Sequence
	14	191	10.0	636	6	BD165437 Novel pol
	15	53	2.8	353	6	AX377473 Sequence
c	16	35	1.8	302070	1	AP005223 Corynebac
	17	23	1.2	121358	9	AF228727 Homo sapi
c	18	23	1.2	179892	2	AC087820 Homo sapi
c	19	23	1.2	183687	9	AC009622 Homo sapi
c	20	22	1.2	161424	2	AC018367 Homo sapi
	21	22	1.2	162495	2	AC015647 Homo sapi
	22	22	1.2	198714	9	CNS05TDW Human chr
	23	22	1.2	202212	9	CNS01RGF Human chr
c	24	21	1.1	68329	2	AC101548 Mus muscu
c	25	21	1.1	149567	10	AC136639 Mus muscu
c	26	21	1.1	154090	2	AC144580 Mus muscu
c	27	21	1.1	156629	2	AC141629 Mus muscu
	28	21	1.1	166631	2	AC131770 Mus muscu
	29	21	1.1	170607	2	AC116502 Mus muscu
c	30	21	1.1	175475	2	AC123631 Mus muscu
c	31	21	1.1	182444	2	AC116481 Mus muscu
c	32	21	1.1	185452	2	AC114566 Mus muscu
	33	21	1.1	187381	2	AC135667 Mus muscu
c	34	21	1.1	190309	10	AC121771 Mus muscu

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 21:26:30 ; Search time 387 Seconds  
(without alignments)  
13267.006 Million cell updates/sec

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcgtccaaggaac 1902

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result			% Query		DB	ID	Description
			Match	Length			
No.	Score						
	1	1902	100.0	1902	24	ABL40476	C. glutamicum luxS
c	2	1902	100.0	309400	22	AAH68534	C glutamicum codin
	3	1311	68.9	1311	22	AAH68204	C glutamicum codin
	4	858	45.1	858	25	ACA01369	C. glutamicum deri
c	5	639	33.6	639	22	AAH68202	C glutamicum codin
	6	411	21.6	426	25	ACA01379	C. glutamicum deri
	7	404	21.2	1052	24	ABA91930	Corynebacterium gl
c	8	319	16.8	588	22	AAH68203	C glutamicum codin
	9	291	15.3	759	22	AAF71286	Corynebacterium gl
c	10	280	14.7	549	25	ACA01390	C. glutamicum deri
	11	191	10.0	636	22	AAH68201	C glutamicum codin
	12	53	2.8	353	24	ABA91931	Corynebacterium gl
	13	20	1.1	20	24	ABL40477	C. glutamicum luxS
c	14	20	1.1	20	24	ABL40478	C. glutamicum luxS
	15	20	1.1	20	24	ABA91932	Corynebacterium gl
c	16	20	1.1	382	14	AAQ59551	Human brain Expres
c	17	20	1.1	1323	24	ABK98540	Botulinum neurotox
c	18	20	1.1	1323	24	ABK98550	Botulinum neurotox
c	19	20	1.1	2475	24	ABK98551	Botulinum neurotox
	20	20	1.1	4827	22	AAI59592	Human polynucleoti
c	21	20	1.1	9486	23	AAS66084	DNA encoding novel
	22	20	1.1	9486	23	AAS90982	DNA encoding novel
c	23	20	1.1	10196	24	AAI64255	Human SLC5A3 gene
c	24	20	1.1	10302	24	AAI72720	BCW2 cDNA. Homo s
c	25	20	1.1	23213	24	ABV78027	Hypoxia-regulated
c	26	19	1.0	383	25	ABX45609	Bovine EST associa
c	27	19	1.0	559	21	AAA31333	Plant microsatelli
c	28	19	1.0	1108	22	AAF71333	Corynebacterium gl
c	29	19	1.0	1116	22	AAH67595	C glutamicum codin
	30	19	1.0	1999	24	ABQ61023	FLJ22833 fis clone
c	31	19	1.0	2811	23	ABL25174	Drosophila melanog
c	32	19	1.0	30390	23	AAS59520	Propionibacterium
	33	19	1.0	349980	22	AAH68532	C glutamicum codin
c	34	18	0.9	264	25	ABX23491	Human GDP-mannose
c	35	18	0.9	490	22	AAK76336	Human immune/haema
c	36	18	0.9	496	22	AAK59840	Human immune/haema
c	37	18	0.9	653	21	AAF12197	Aspergillus oryzae
	38	18	0.9	658	24	ABX65502	Helicobacter pylor
	39	18	0.9	672	23	ABL13199	Drosophila melanog
	40	18	0.9	723	22	AAH67475	C glutamicum codin
	41	18	0.9	723	25	ACA01224	C. glutamicum deri
c	42	18	0.9	969	25	ACA01000	C. glutamicum deri
c	43	18	0.9	978	22	AAH66424	C glutamicum codin
	44	18	0.9	1001	21	AAC57967	Arachidonic acid m
	45	18	0.9	1021	24	ABX66099	Helicobacter pylor

# ALIGNMENTS



## SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	1902	100.0	1902	10	US-09-824-551-1	Sequence 1, Appli
c	2	1902	100.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
	3	1311	68.9	1311	10	US-09-738-626-3239	Sequence 3239, Ap
c	4	639	33.6	639	10	US-09-738-626-3237	Sequence 3237, Ap
c	5	319	16.8	588	10	US-09-738-626-3238	Sequence 3238, Ap
	6	191	10.0	636	10	US-09-738-626-3236	Sequence 3236, Ap
	7	22	1.2	516	13	US-10-027-632-287113	Sequence 287113,
c	8	21	1.1	3138	14	US-10-156-761-2538	Sequence 2538, Ap
	9	21	1.1	9025608	14	US-10-156-761-1	Sequence 1, Appli
	10	20	1.1	20	10	US-09-824-551-3	Sequence 3, Appli
c	11	20	1.1	20	10	US-09-824-551-4	Sequence 4, Appli
c	12	20	1.1	1323	13	US-10-011-588-10	Sequence 10, Appl
c	13	20	1.1	1323	13	US-10-011-588-30	Sequence 30, Appl
c	14	20	1.1	2475	13	US-10-011-588-32	Sequence 32, Appl
c	15	19	1.0	383	10	US-09-960-352-10774	Sequence 10774, A
c	16	19	1.0	611	13	US-10-027-632-192050	Sequence 192050,
c	17	19	1.0	961	13	US-10-027-632-251784	Sequence 251784,
c	18	19	1.0	1116	10	US-09-738-626-2630	Sequence 2630, Ap
	19	19	1.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
c	20	18	0.9	261	14	US-10-156-761-494	Sequence 494, App
c	21	18	0.9	264	10	US-09-878-574-5550	Sequence 5550, Ap
c	22	18	0.9	461	14	US-10-171-581-314	Sequence 314, App
c	23	18	0.9	536	13	US-10-027-632-136635	Sequence 136635,
c	24	18	0.9	687	13	US-10-027-632-106927	Sequence 106927,
c	25	18	0.9	687	13	US-10-027-632-106928	Sequence 106928,
c	26	18	0.9	687	13	US-10-027-632-136634	Sequence 136634,
	27	18	0.9	723	10	US-09-738-626-2510	Sequence 2510, Ap
c	28	18	0.9	727	13	US-10-027-632-25349	Sequence 25349, A
c	29	18	0.9	727	13	US-10-027-632-25350	Sequence 25350, A
	30	18	0.9	800	13	US-10-027-632-154095	Sequence 154095,
c	31	18	0.9	931	12	US-10-017-161-1863	Sequence 1863, Ap
c	32	18	0.9	978	10	US-09-738-626-1459	Sequence 1459, Ap
c	33	18	0.9	1079	14	US-10-078-090-3	Sequence 3, Appli
	34	18	0.9	1468	12	US-09-882-227-259	Sequence 259, App
	35	18	0.9	2007	10	US-09-801-368-237	Sequence 237, App
c	36	18	0.9	3393	14	US-10-128-714-203	Sequence 203, App
	37	18	0.9	3401	14	US-10-081-119-19	Sequence 19, Appl
c	38	18	0.9	3867	14	US-10-128-714-202	Sequence 202, App
c	39	18	0.9	4035	14	US-10-128-714-5202	Sequence 5202, Ap
	40	18	0.9	13584	12	US-10-311-455-587	Sequence 587, App
	41	18	0.9	536165	11	US-09-939-964-1	Sequence 1, Appli
c	42	18	0.9	1830121	14	US-10-329-960-1	Sequence 1, Appli
c	43	18	0.9	9025608	14	US-10-156-761-1	Sequence 1, Appli
	44	17	0.9	160	10	US-09-783-590-8947	Sequence 8947, Ap
	45	17	0.9	186	10	US-09-878-574-3194	Sequence 3194, Ap

## ALIGNMENTS

RESULT 1  
US-09-824-551-1





28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%		Query		DB	ID	Description
			Match	Length			
No.	Score						
1	23	1.2	383	10	BF910515	BF910515	IL2-UT007
c 2	22	1.2	591	29	BZ492878	BZ492878	BONBP91TF
c 3	22	1.2	856	29	BZ497273	BZ497273	BONSE69TR
4	21	1.1	345	28	BH265149	BH265149	CH230-55I
c 5	21	1.1	642	29	BX192671	BX192671	Danio rer
6	21	1.1	767	28	AZ717195	AZ717195	RPCI-24-1
7	21	1.1	805	29	BZ209421	BZ209421	CH230-342
8	20	1.1	192	28	BH174645	BH174645	Bir49L18
c 9	20	1.1	273	10	BF716978	BF716978	NcEST3a19
c 10	20	1.1	273	14	N88922	N88922	K6584F Huma
11	20	1.1	292	28	AZ007734	AZ007734	RPCI-23-2
c 12	20	1.1	382	14	M78393	M78393	EST00541 Fe
13	20	1.1	394	12	BI058627	BI058627	CM3-GN029
c 14	20	1.1	485	12	BI515786	BI515786	BB160020B
c 15	20	1.1	486	28	AZ009964	AZ009964	RPCI-23-2
16	20	1.1	659	14	CB267786	CB267786	1006692 H
c 17	20	1.1	746	29	AG134951	AG134951	Pan trogl
c 18	20	1.1	859	29	CC400720	CC400720	PUHLM18TD
c 19	20	1.1	936	29	BZ571963	BZ571963	msh2_2170
c 20	20	1.1	1058	12	BM562952	BM562952	AGENCOURT
c 21	20	1.1	1543	29	CC208526	CC208526	CH261-108
c 22	19	1.0	256	12	BI293206	BI293206	UI-R-DK0-
23	19	1.0	331	9	AA375611	AA375611	EST88182
c 24	19	1.0	359	9	AL968593	AL968593	AL968593
25	19	1.0	368	9	AA781849	AA781849	ai61g09.s
c 26	19	1.0	414	9	AV690055	AV690055	AV690055
27	19	1.0	416	13	BY208181	BY208181	BY208181
28	19	1.0	418	10	BE493362	BE493362	WHE0571_H
29	19	1.0	422	6	AL812086	AL812086	Triticum
c 30	19	1.0	427	13	BQ529073	BQ529073	3524_1_44
31	19	1.0	432	13	BY281892	BY281892	BY281892
c 32	19	1.0	436	9	AL786619	AL786619	AL786619
33	19	1.0	457	13	BY561821	BY561821	BY561821
34	19	1.0	458	10	BF717012	BF717012	NcEST3a19
35	19	1.0	459	12	BM147775	BM147775	TCAAP1Q14
c 36	19	1.0	465	9	AL645006	AL645006	AL645006
c 37	19	1.0	467	13	BX088966	BX088966	BX088966
38	19	1.0	476	9	AA147562	AA147562	z151d01.r
c 39	19	1.0	483	14	CA826740	CA826740	1114004C1
c 40	19	1.0	495	13	BQ442032	BQ442032	3524_1_7
41	19	1.0	503	9	AV604523	AV604523	AV604523
c 42	19	1.0	503	13	BU097610	BU097610	3524_1_19
c 43	19	1.0	505	9	AW418953	AW418953	ha24d06.x
c 44	19	1.0	507	9	AL801184	AL801184	AL801184
c 45	19	1.0	508	9	AL970492	AL970492	AL970492

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 22:41:31 ; Search time 102 Seconds  
(without alignments)  
8230.497 Million cell updates/sec

Title: US-09-824-551-1  
Perfect score: 1902  
Sequence: 1 ggtaggagtaaaaaacgcag.....ggccccggcgccaaggaac 1902

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%					Description
	Query	Match	Length	ID		
No.	Score					
	1	18	0.9	657	4	US-09-252-991A-7507
	2	18	0.9	1001	4	US-09-641-638-601
c	3	18	0.9	1185	4	US-09-252-991A-1931
	4	18	0.9	1287	4	US-09-252-991A-7363
c	5	18	0.9	1392	4	US-09-252-991A-7136
c	6	18	0.9	2163	3	US-09-296-284-22
c	7	18	0.9	2232	5	PCT-US96-05320A-113
c	8	18	0.9	2265	3	US-09-296-284-1
c	9	18	0.9	4830	3	US-09-296-284-7
	10	18	0.9	536165	4	US-09-214-808-1
c	11	18	0.9	1830121	4	US-09-557-884-1